Methods for variable time-delay estimation in industrial data

ENBIS, Trondheim 26-29 June 2022

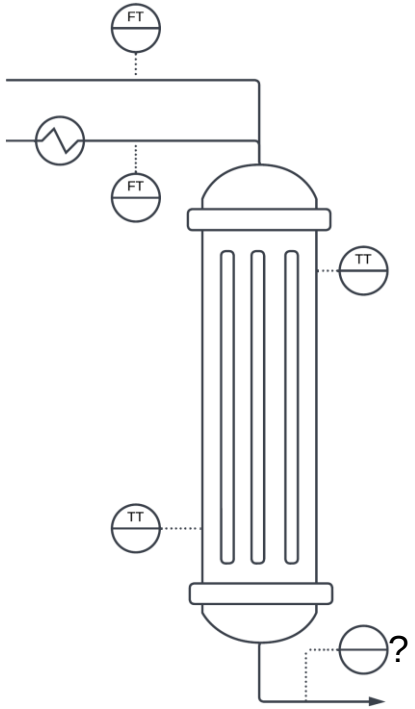
Marco Cattaldo, Alberto Ferrer, Ingrid Måge

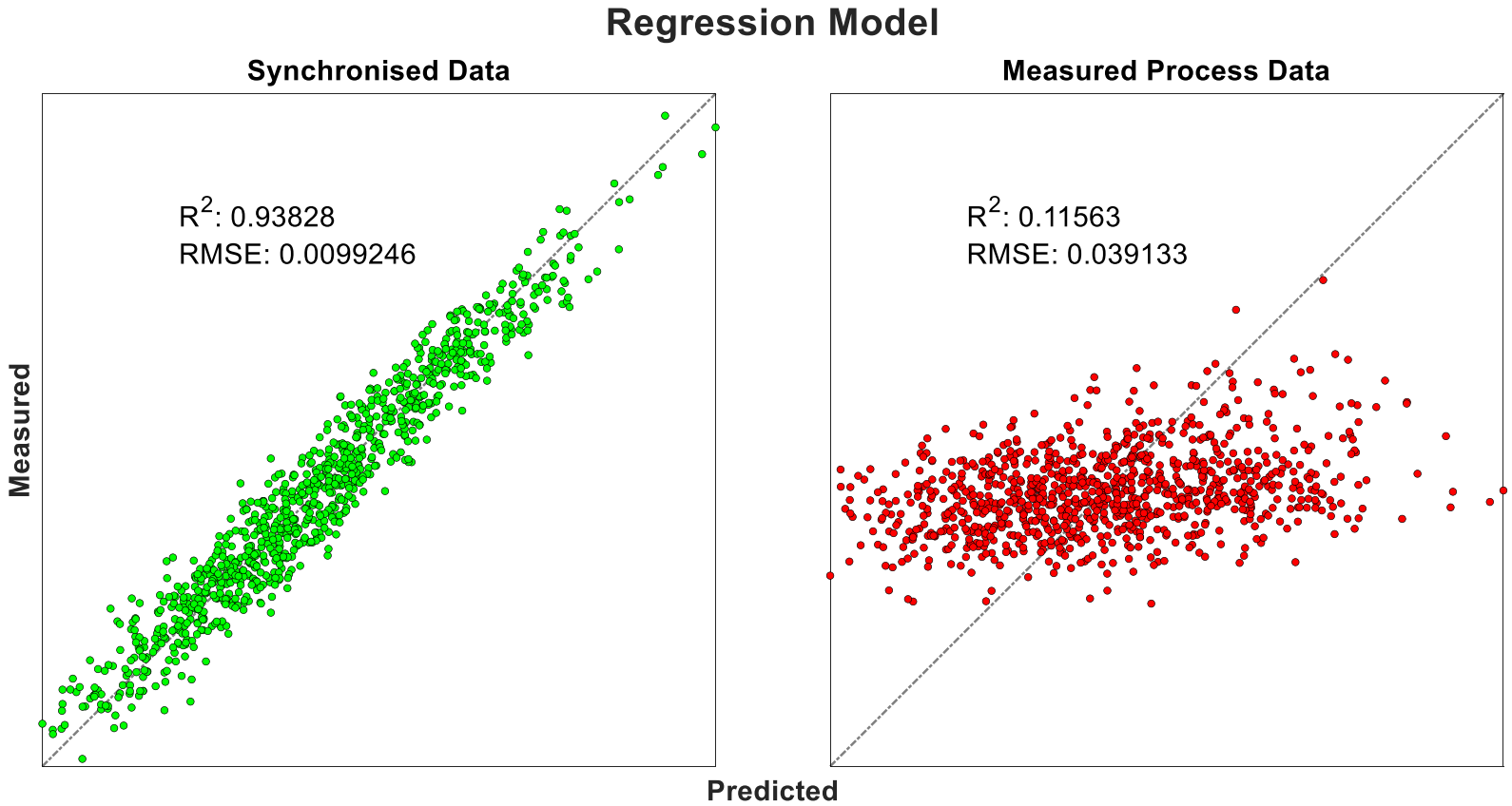
Nofima
Norwegian Institute of Food, Fisheries and Aquaculture Research

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

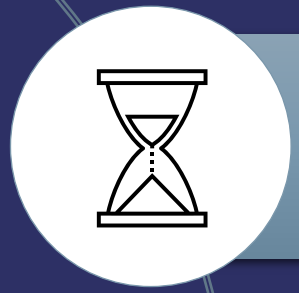
Have you ever…

Not so imaginary data story





Presentation Overview



What is variable time-delay

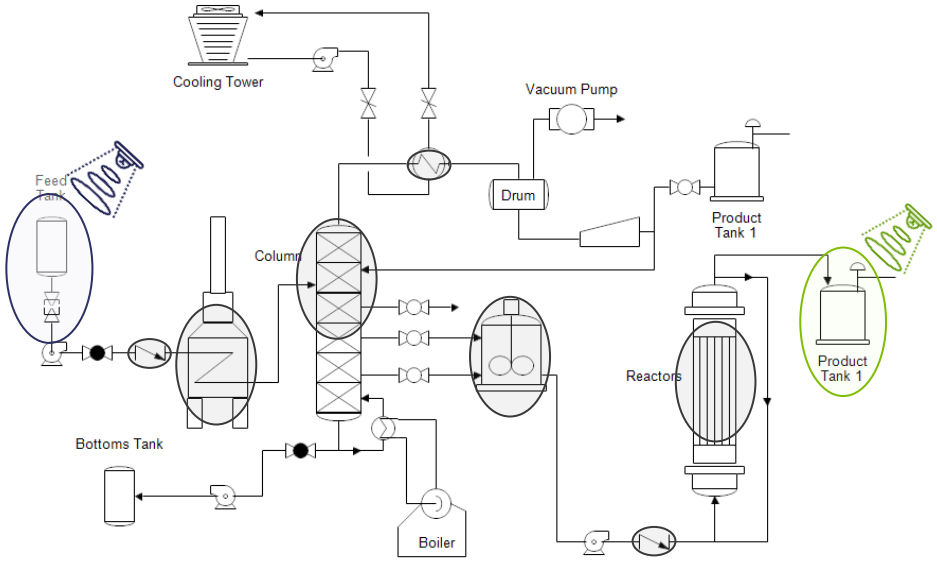
How do you estimate variable time-delay

Use Case

Variable time-delay estimation
Synchronisation of process variables



$$\textbf{\textit{Product quality}} = function(\text{raw material}, \text{processing})$$



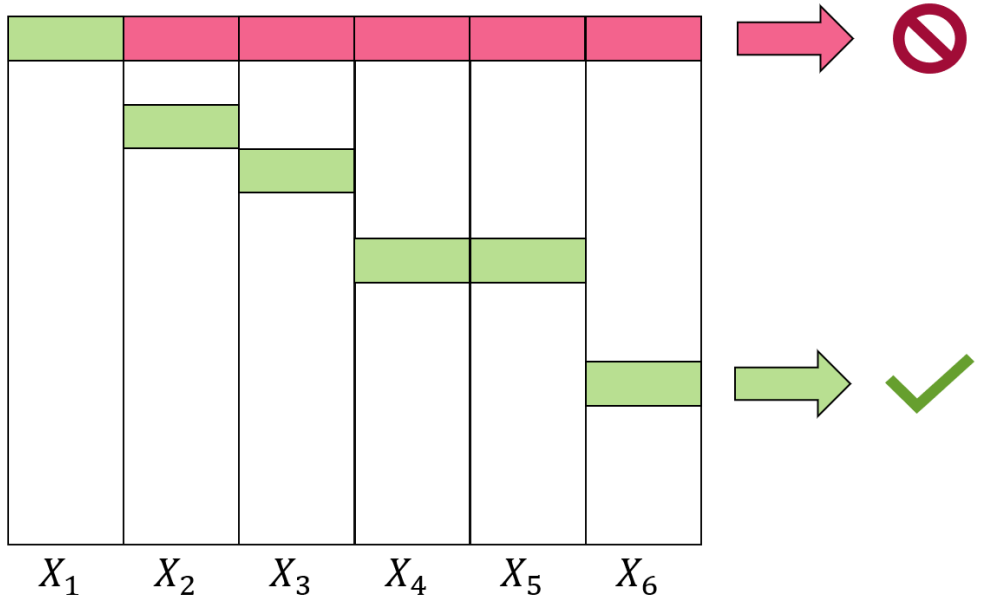
- Variable time delay is structural to most *continuous* processes.

- Data acquisition is *simultaneous*, process flow is *not*

Variable time-delay estimation
Approaching the time-delay estimation problem (1/2)

- Not tackling the issue will give *wrong* results

- How to solve the Variable time delay estimation problem?

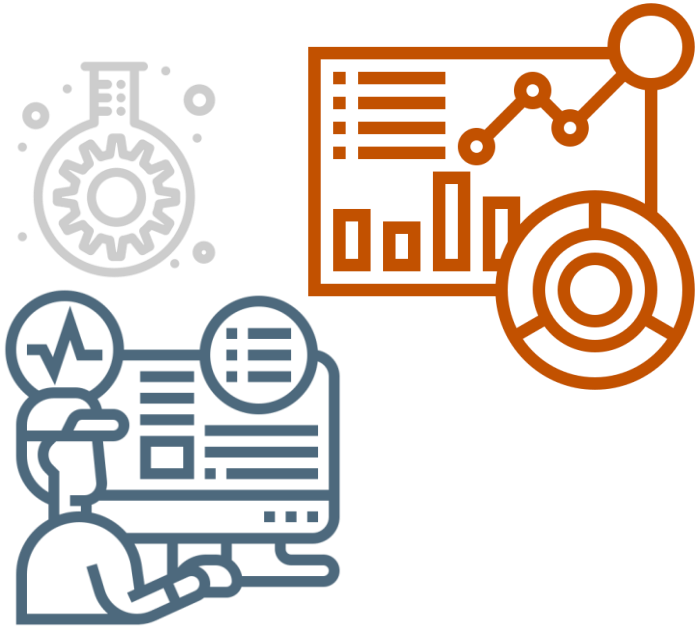


Variable time-delay estimation
Approaching the time-delay estimation problem (2/2)

- *Physical* variable time-delay estimation
  - Serious feasibility issues

- Consulting the *process operators*
  - Not precise enough

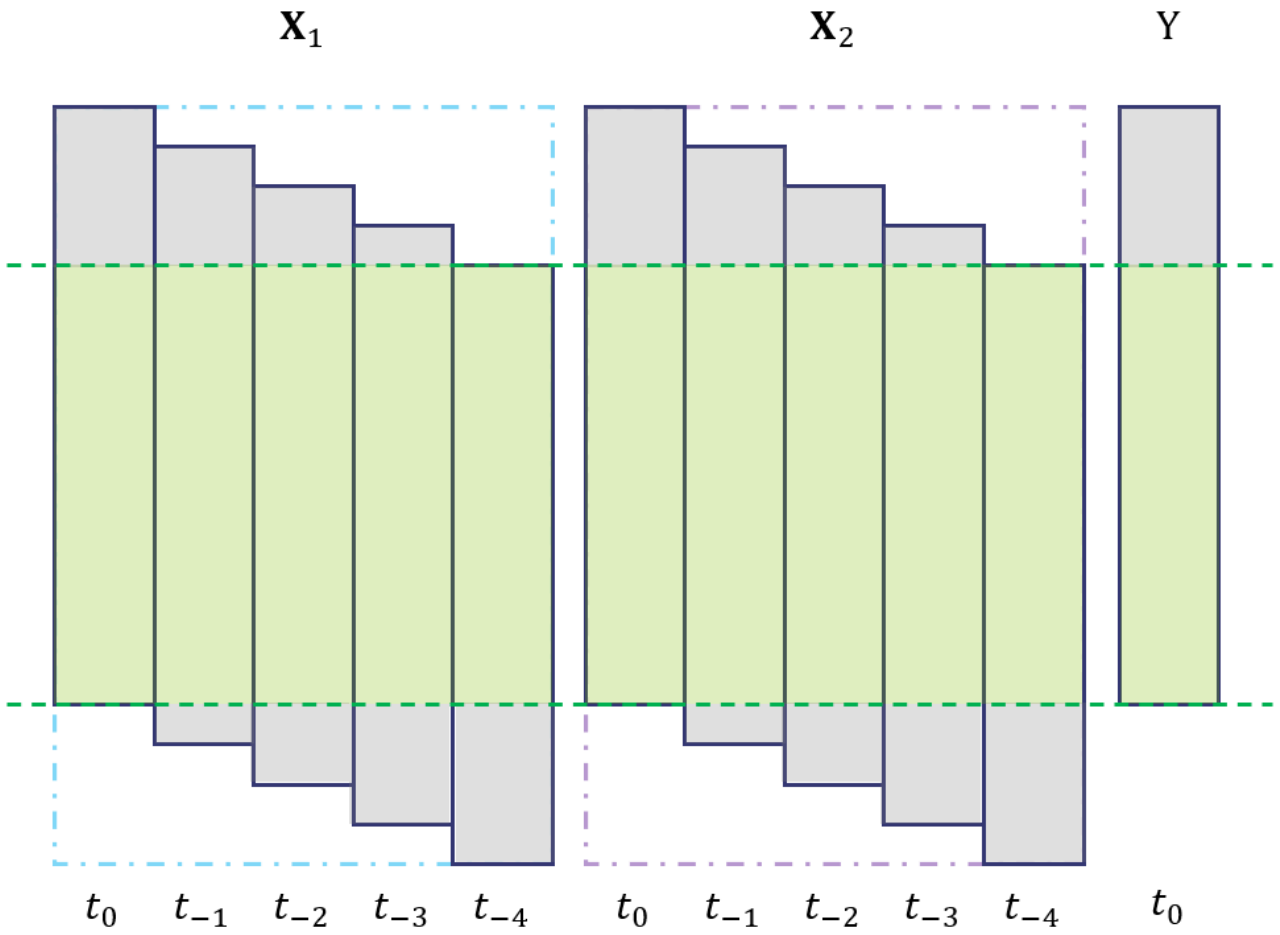
- A large amount of data-driven methods are available



Data-driven variable time-delay estimation how to:
Step 1: Incorporate all lags (1/3)

Steps for delay estimation

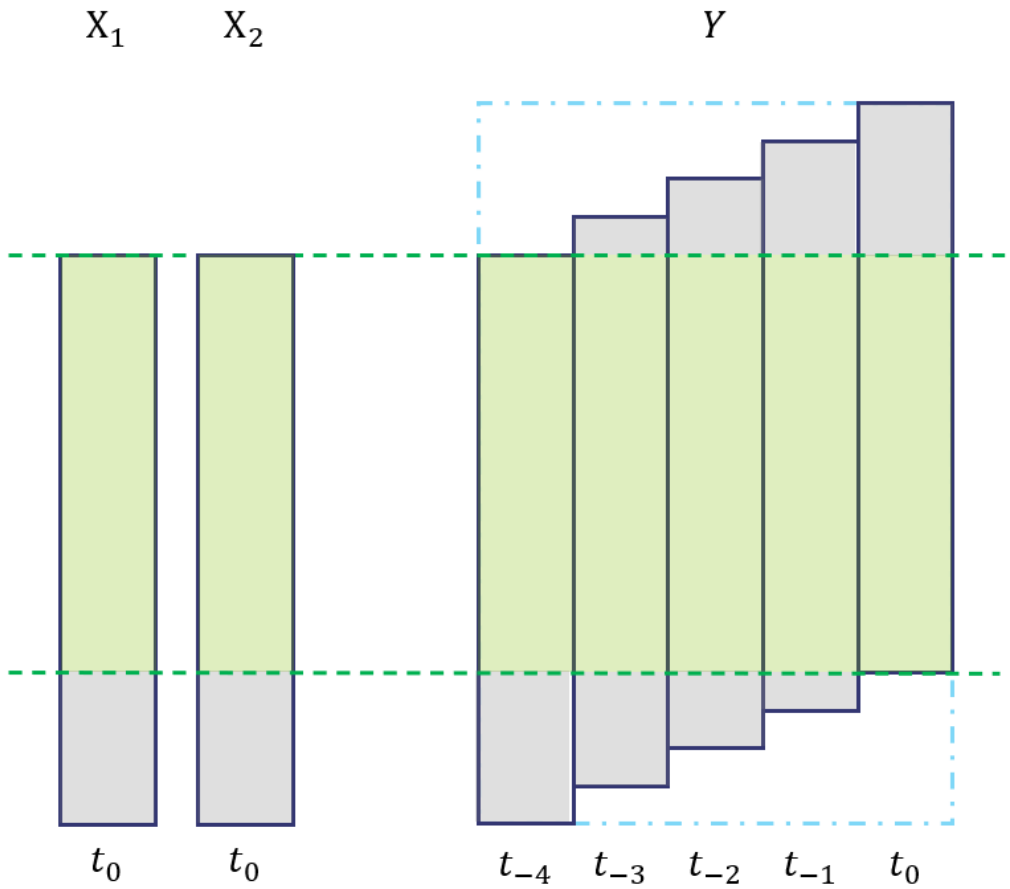
- Include all Lags in the dataset
  - Lag augmented predictor matrix
  - Lag augmented target matrix
  - Lag selection



Data-driven variable time-delay estimation how to:
Step 1: Incorporate all lags (2/3)

Steps for delay estimation

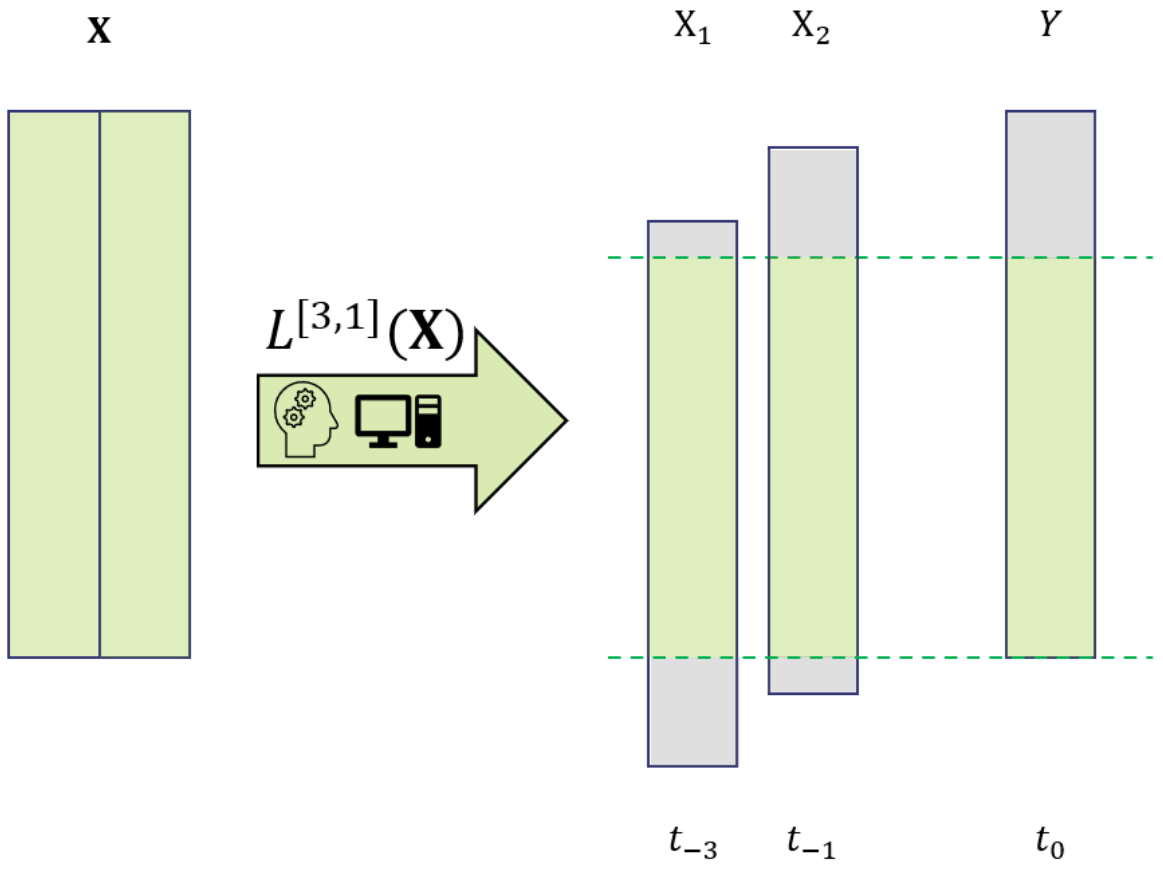
- Include all Lags in the dataset
  - Lag augmented predictor matrix
  - Lag augmented target matrix
  - Lag selection



Data-driven variable time-delay estimation how to:
Step 1: Incorporate all lags (3/3)

Steps for delay estimation

- Include all Lags in the dataset
  - Lag augmented predictor matrix
  - Lag augmented target matrix
  - Lag selection

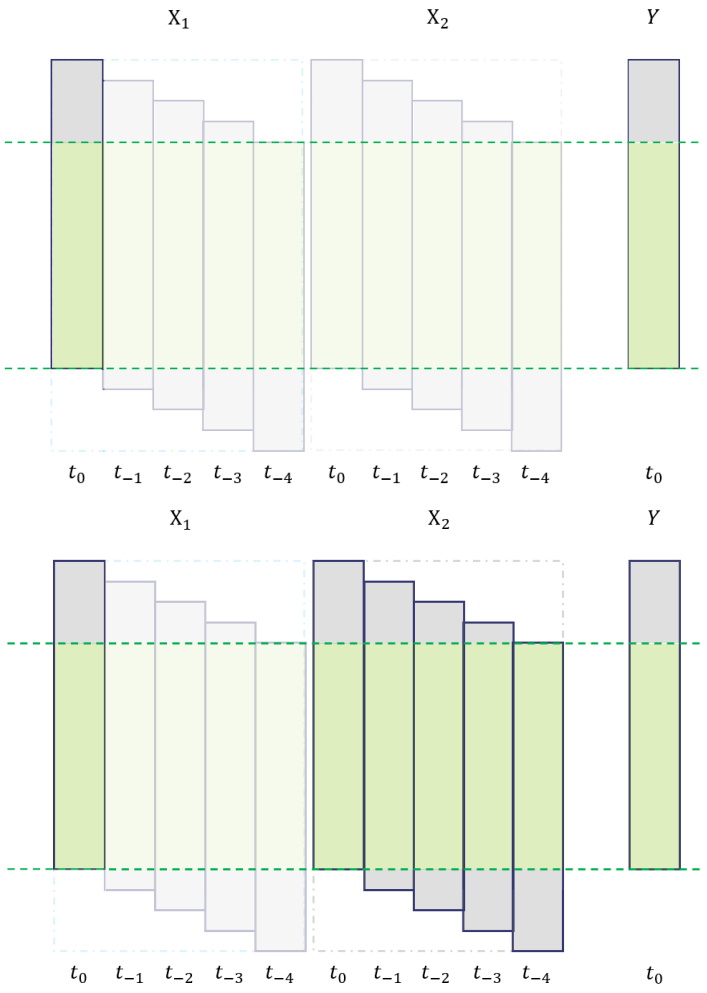


Data-driven variable time-delay estimation how to
Step 2: Identify the correct lags

Steps for delay estimation

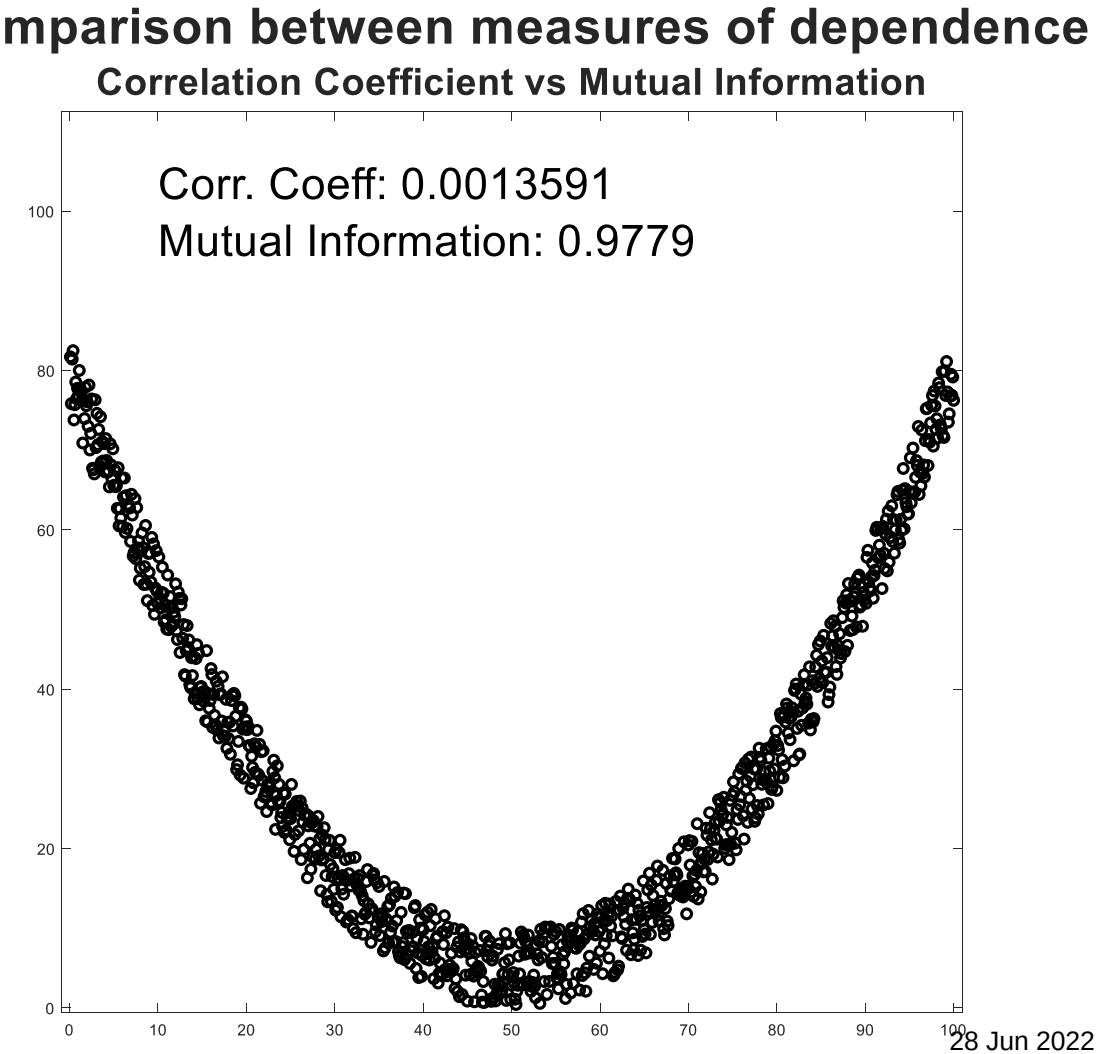
- Two possible approaches
  - Bivariate
  - Multivariate

- In all cases we can use
  - a MODEL
  - an INDEX



Data-driven variable time-delay estimation how to:
A word on Indices and Models

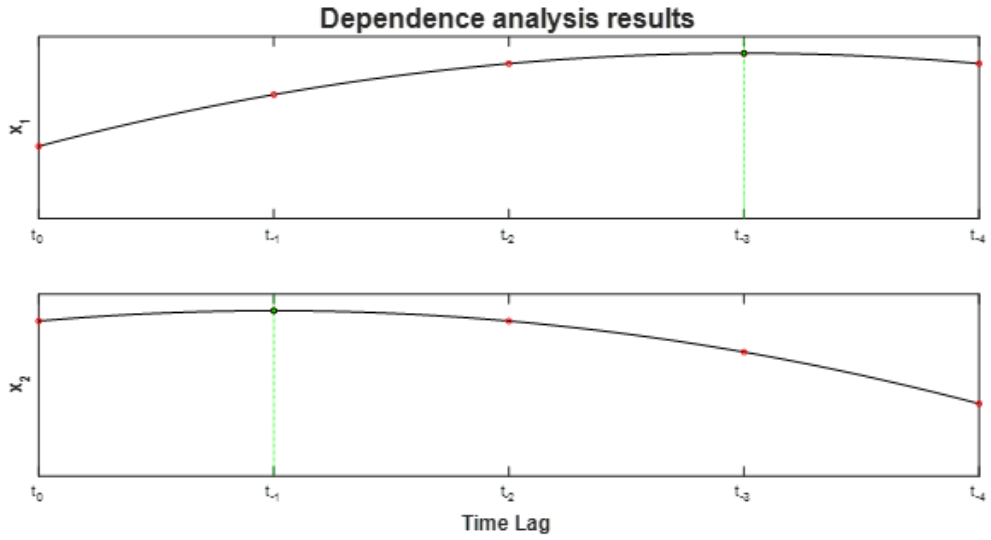
- Beware of nonlinearities!
- Identify the right type of model / index
- Compare different models / Indices

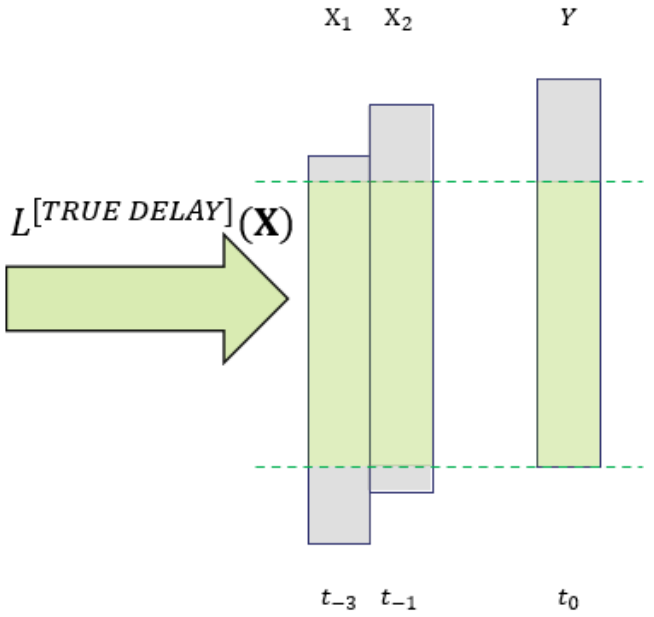


Data-driven variable time-delay estimation how to:
Step 3: Analyse the results

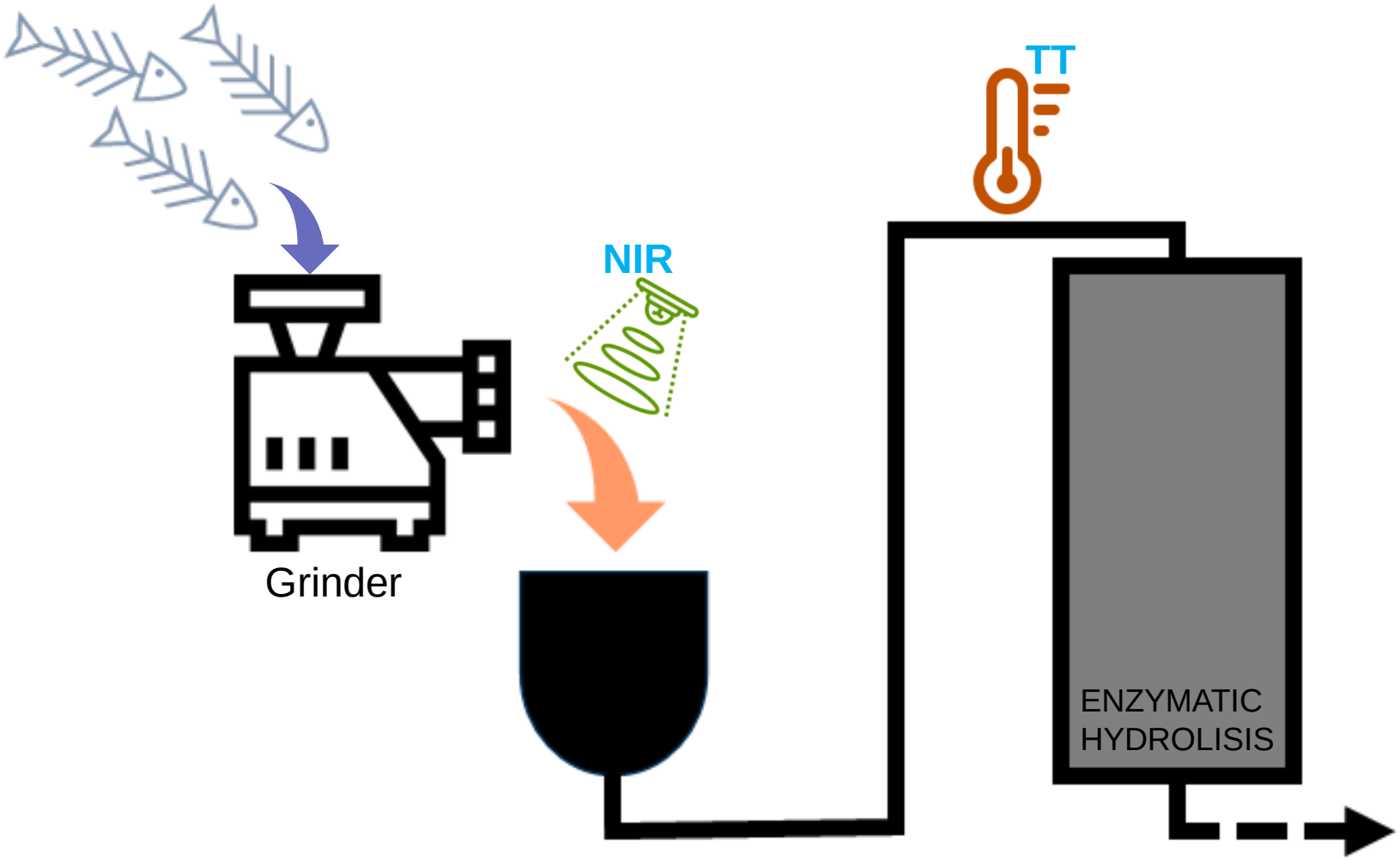
Steps for delay estimation

- Postprocess the results:
  - Identify max lag(s)
  - Rebuild the X matrix incorporating only the correct lags
- Proceed with further analysis



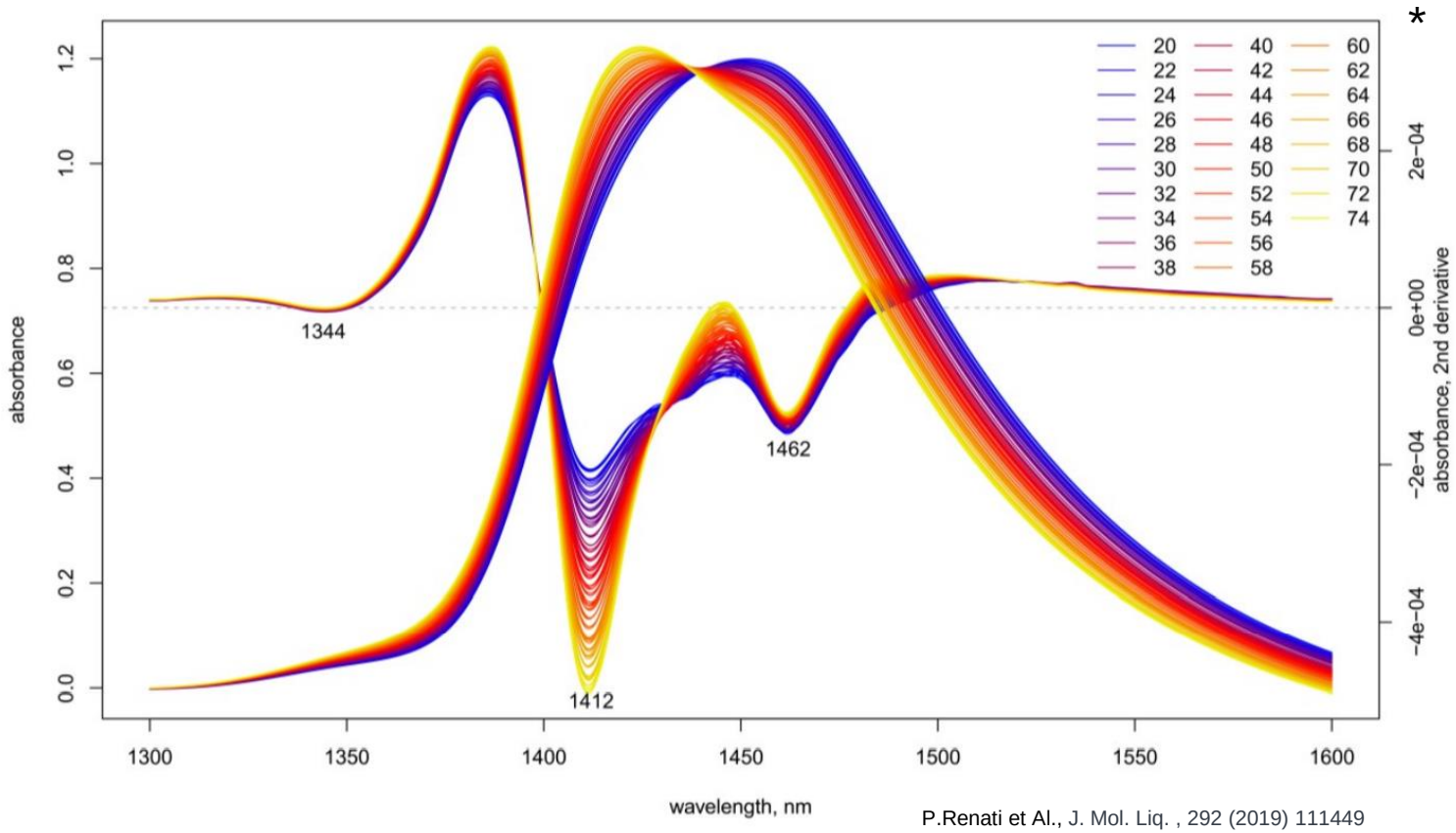


Industrial case:
In-Line NIR system @ Biomega



Industrial case:
Temperature and NIR (1/2)

- NIR measures light absorption on a sample
- With temperature the absorption changes
- NIR is an expensive thermometer

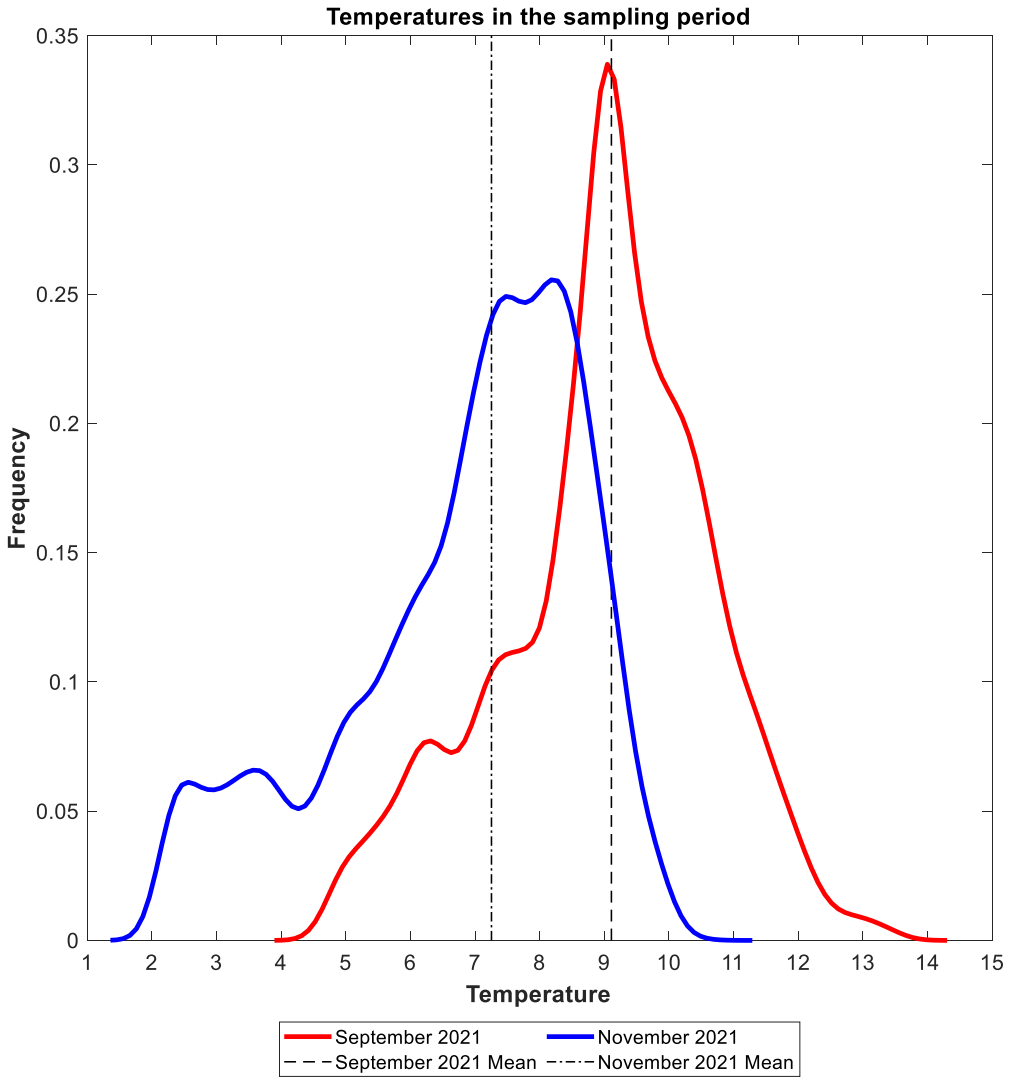


P.Renati et Al., J. Mol. Liq. , 292 (2019) 111449

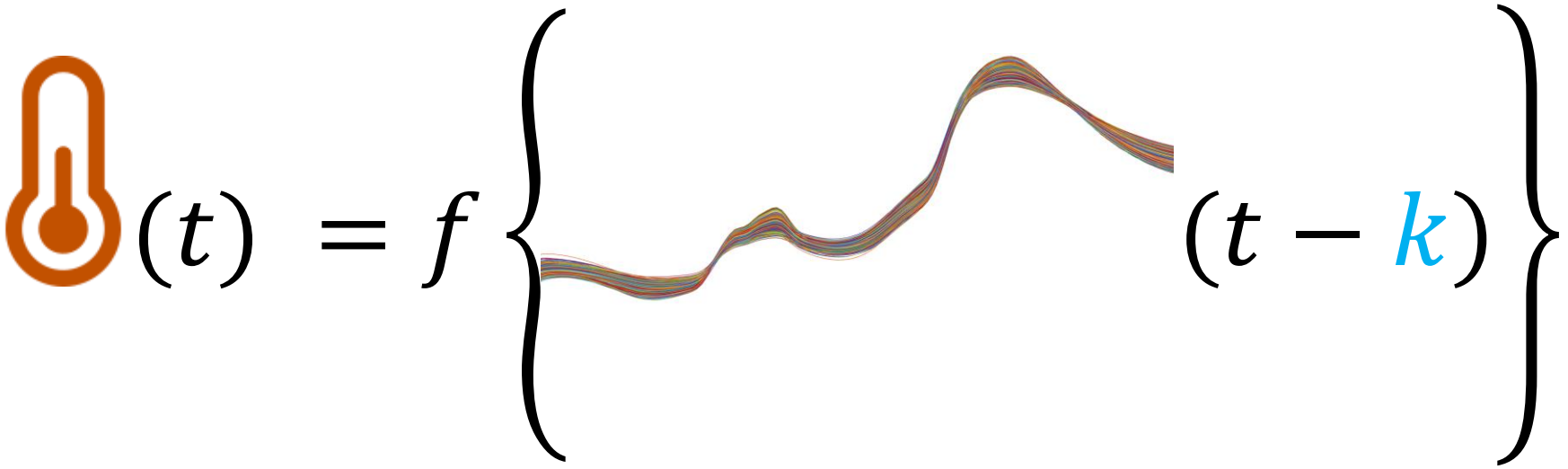
Industrial case:
Temperature and NIR (2/2)

- Great range of temperatures

- Calibration could be made more robust



Industrial case:
Solving the Time delay estimation problem



- For which *k* the relationship is the *strongest?*

- Compare four techniques

Industrial case:
Choice of techniques

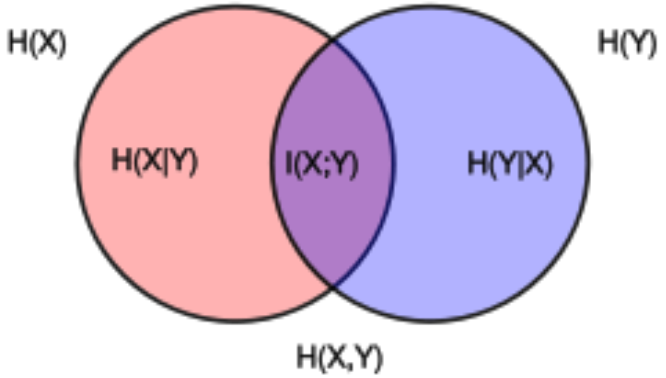
	INDEX	MODEL
Augmented matrix	Correlation Coefficient	PLS – Selectivity Ratio
Lag Selection	Mutual Information	PLS – Goodness of Fit

- Different kind of approaches tested
  - Bivariate
  - Multivariate

- Different parameters tested
  - Mutual Information, *number of neighbors*
  - PLS, *number of components*

Industrial case:
Techniques Explanation

- Mutual Information – I(**X**, Y)
  - Decrease in the uncertainty (entropy) of Y by knowing **X**
  - Distance between $P(\boldsymbol{X}, Y)$ and $P(\boldsymbol{X})P(Y)$



- Zero iif **X** and Y are independent

- Concept formulated for discrete signals
  - Must be *estimated* for continuous signals

Industrial case:
Interpreting the results

	t_0	t_{10}	t_{20}	t_{30}	t_{40}	t_{50}	t_{60}	t_{70}	t_{80}	t_{90}	t_{100}	t_{110}	t_{120}	t_{130}	t_{140}	t_{150}	t_{160}	t_{170}	t_{180}	t_{190}	t_{200}
Correlation Coefficient																					+
PLS – Selectivity Ratio									+						++					++	
PLS – R^2 of Cross Validation															+	+	++	+			
Mutual Information																+++	++				
Total									+						+++	++++	++++	+		++	+

- Some techniques are less robust

- The expected delay between NIR and Temperature is 140/160 seconds

Variable Time Delay estimation:
take-home points

- Variable Time-delay needs to be accounted for

- Many available methods

- No Free Lunch
  - there are *no catch-all methods*

Nofima
Norwegian Institute of Food, Fisheries and Aquaculture Research

Thanks for your attention!

Time Delay Estimation

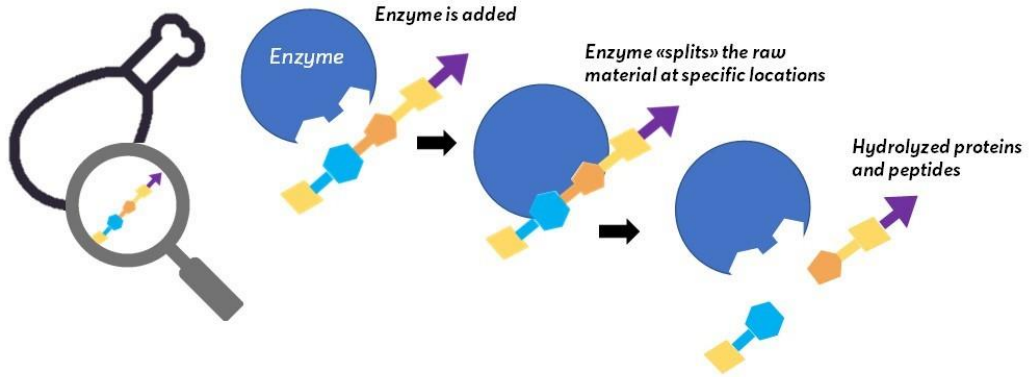
Feature Selection and Time Delay Estimation

- Feature selection and time delay estimation have a lot of common ground
  - Time delay estimation could be seen as a particular case Feature Selection on the lagged matrix

- Characteristics of time delay estimation
  1. Known number of variables to search for
  2. Same process generates all the lagged versions of a variable
     - There should be a clear maximum point in the dependence measurements vector
  3. The number of variables to target dramatically increases compared to the original Feature selection problem
     - l possible states for m variables→ $C^R(\mathrm{m},\mathrm{l}) = \frac{(m+l-1)!}{m!(l-1)!}$
     - Search strategies need to be further optimized for the methods to be applied

Industrial case:
In-Line NIR system @ Biomega

- Innovative company
- Extracts protein from fish rest raw material
- Sells to Human and animal feed market



Extra Slides End